

SEQUENCE LISTING

<110> Bayer Aktiengesellschaft

<120> Nucleic acids which encode
insect acetylcholine receptor subunits

<130> Le A 33 020-Foreign Countries

<140>

<141>

<150> DE-198 19 829.9

<151> 1998-05-04

<160> 6

<170> PatentIn Ver. 2.1

<210> 1

<211> 2886

<212> DNA

<213> Drosophila melanogaster

<220>

<221> CDS

<222> (372) .. (2681)

<400> 1

ggcagcagaa aaagttgtgg tataaacttt tattgtagga aaacgcataa aaataataga 60

aaaacgctct tcgggttgta aagaaaataa gaagacaaaa gaaagacatg aaaacgttgc 120

aaacaataaa gcatatactt gccatattga tataaagggg aatcgtgaaa aggcggtgaa 180

aatttcgtaa gattagttgg tattaagggc agcccatgca cacagctaaa aagggaacta 240

aaaaaacccc gcacagaaca atgaaagctg cagcagctgg ataaggccga caaaccgaa 300

aattatatta ttgtaatcta gtagagagca gacaacatat ccgctggcaa caaccaacac 360

cgaaagagac t atg aaa aat gca caa ctg aaa ctg act gaa gtt gac gat 410

Met Lys Asn Ala Gln Leu Lys Leu Thr Glu Val Asp Asp

1

5

10

gat gag ctg tgg ctg gca gta aga tta gcg cac tgc agc agc aac ttt 458

Asp Glu Leu Trp Leu Ala Val Arg Leu Ala His Cys Ser Ser Asn Phe

15

20

25

agc agc agt agc agc aca aga acc acc agc agc aac cag agg cac aac 506

Ser Ser Ser Ser Ser Thr Arg Thr Thr Ser Ser Asn Gln Arg His Asn

30

35

40

45

cag caa ctc aca aca ctg caa cca agg agc tta agt aca aaa cac cac 554

Gln Gln Leu Thr Thr Leu Gln Pro Arg Ser Leu Ser Thr Lys His His

50

55

60

agc aac att gca agc gag cag cac aat agc cag caa cag gag cca gca 602

Ser Asn Ile Ala Ser Glu Gln His Asn Ser Gln Gln Gln Glu Pro Ala

65

70

75

tcg aag gac gag gat gta gcc aac cac ggt aga agc aat gac cag cag 650

Ser Lys Asp Glu Asp Val Ala Asn His Gly Arg Ser Asn Asp Gln Gln

80 85 90

acg cat ctg caa cag cta gac agc agc aac atg ttg tgc cca aag aca 698
 Thr His Leu Gln Gln Leu Asp Ser Ser Asn Met Leu Ser Pro Lys Thr
 95 100 105

gcc gca gca gca act gct gcc ggc gat gaa gca aca acc caa caa cca 746
 Ala Ala Ala Ala Thr Ala Ala Gly Asp Glu Ala Thr Thr Gln Gln Pro
 110 115 120 125

aca aac ata aga ctg tgt gca cgc aag cga caa cga ttg cgt cgc cga 794
 Thr Asn Ile Arg Leu Cys Ala Arg Lys Arg Gln Arg Leu Arg Arg Arg
 130 135 140

cga aaa aga aaa cca gca acc cca aac gaa aca gat atc aag aaa caa 842
 Arg Lys Arg Lys Pro Ala Thr Pro Asn Glu Thr Asp Ile Lys Lys Gln
 145 150 155

cag caa ctt agc atg cct ccc ttc aaa acg cgc aaa tcc acg gac acc 890
 Gln Gln Leu Ser Met Pro Pro Phe Lys Thr Arg Lys Ser Thr Asp Thr
 160 165 170

tac agc aca cca gca gca aca acc agc tgt ccg aca gcc acc tac atg 938
 Tyr Ser Thr Pro Ala Ala Thr Thr Ser Cys Pro Thr Ala Thr Tyr Met
 175 180 185

caa tgt cga gcc agc gac aat gag ttc agt att ccg ata tgc aga cat 986
 Gln Cys Arg Ala Ser Asp Asn Glu Phe Ser Ile Pro Ile Ser Arg His
 190 195 200 205

gat aga gta tcc acg gcc aca ttc gcc tgg gtg ttg cat gtg ctg cag 1034

Asp Arg Val Ser Thr Ala Thr Phe Ala Trp Val Leu His Val Leu Gln

210

215

220

gtg ctg ctc gtg tcg ctg caa cag tgg caa ctt cac gtg caa cag cga 1082

Val Leu Leu Val Ser Leu Gln Gln Trp Gln Leu His Val Gln Gln Arg

225

230

235

tcg gtg cta ctg ttc aga agg atc gca gcg agc acc atc gcc ttc att 1130

Ser Val Leu Leu Phe Arg Arg Ile Ala Ala Ser Thr Ile Ala Phe Ile

240

245

250

tcc tat tta ggc agc ttt gca gcg caa ctg aaa aat agc agc agc agc 1178

Ser Tyr Leu Gly Ser Phe Ala Ala Gln Leu Lys Asn Ser Ser Ser Ser

255

260

265

agt agc agc agc aac agc agc aac aac agc agc acg caa ata tta aac 1226

Ser Ser Ser Ser Asn Ser Ser Asn Asn Ser Ser Thr Gln Ile Leu Asn

270

275

280

285

gga ctt aat aaa cac tca tgg ata ttt tta ttg ata tat ttg aat tta 1274

Gly Leu Asn Lys His Ser Trp Ile Phe Leu Leu Ile Tyr Leu Asn Leu

290

295

300

tct gct aaa gtt tgc cta gca gga tat cat gaa aag aga ctg tta cac 1322

Ser Ala Lys Val Cys Leu Ala Gly Tyr His Glu Lys Arg Leu Leu His

305

310

315

gat ctt ttg gat cct tat aat aca cta gaa cgt ccc gtt ctc aat gaa 1370

Asp Leu Leu Asp Pro Tyr Asn Thr Leu Glu Arg Pro Val Leu Asn Glu

320

325

330

tcg gac ccg tta caa tta agc ttt ggt tta act tta atg caa att atc 1418
 Ser Asp Pro Leu Gln Leu Ser Phe Gly Leu Thr Leu Met Gln Ile Ile

335

340

345

gat gtg gac gag aaa aat caa ttg cta gtc act aat gtg tgg tta aaa 1466
 Asp Val Asp Glu Lys Asn Gln Leu Leu Val Thr Asn Val Trp Leu Lys

350

355

360

365

ctg gag tgg aac gac atg aat ctc cgc tgg aac acc ttc gac tat ggc 1514
 Leu Glu Trp Asn Asp Met Asn Leu Arg Trp Asn Thr Ser Asp Tyr Gly

370

375

380

gga gtt aag gat ctg cga ata ccg ccg cat cgc atc tgg aag ccg gac 1562
 Gly Val Lys Asp Leu Arg Ile Pro Pro His Arg Ile Trp Lys Pro Asp

385

390

395

gtg ctg atg tac aac agt gcg gat gag gga ttt gac ggc acc tac cag 1610
 Val Leu Met Tyr Asn Ser Ala Asp Glu Gly Phe Asp Gly Thr Tyr Gln

400

405

410

acg aac gtg gtg gtg cgg aac aac ggc tcg tgt cta tac gtt ccg ccg 1658
 Thr Asn Val Val Val Arg Asn Asn Gly Ser Cys Leu Tyr Val Pro Pro

415

420

425

ggg atc ttc aag tcg acg tgc aag atc gac atc acg tgg ttc ccc ttc 1706
 Gly Ile Phe Lys Ser Thr Cys Lys Ile Asp Ile Thr Trp Phe Pro Phe

430

435

440

445

gat gac cag cgg tgc gag atg aag ttc ggc agt tgg acc tac gac gga 1754
 Asp Asp Gln Arg Cys Glu Met Lys Phe Gly Ser Trp Thr Tyr Asp Gly

450

455

460

ttc cag ctg gat tta caa tta caa gat gaa act ggc ggt gat atc agc 1802

Phe Gln Leu Asp Leu Gln Leu Gln Asp Glu Thr Gly Gly Asp Ile Ser

465

470

475

agt tac gtg ctc aac ggc gag tgg gaa cta ctg ggt gtg ccc ggc aaa 1850

Ser Tyr Val Leu Asn Gly Glu Trp Glu Leu Leu Gly Val Pro Gly Lys

480

485

490

cgt aac gag atc tat tac aac tgc tgc ccg gaa ccc tat ata gac atc 1898

Arg Asn Glu Ile Tyr Tyr Asn Cys Cys Pro Glu Pro Tyr Ile Asp Ile

495

500

505

acc ttc gcc atc atc atc cgc cga cga aca ctg tac tat ttc ttc aac 1946

Thr Phe Ala Ile Ile Ile Arg Arg Arg Thr Leu Tyr Tyr Phe Phe Asn

510

515

520

525

ctg atc ata cct tgt gta ctg att gcc tcc atg gcc ttg ctc gga ttc 1994

Leu Ile Ile Pro Cys Val Leu Ile Ala Ser Met Ala Leu Leu Gly Phe

530

535

540

acc ctg ccg cca gat tcc ggt gaa aaa tta tcc ctg ggt gtt acc atc 2042

Thr Leu Pro Pro Asp Ser Gly Glu Lys Leu Ser Leu Gly Val Thr Ile

545

550

555

ttg ctc tcc ctg acc gtg ttt ctg aat atg gtt gcc gag aca atg ccg 2090

Leu Leu Ser Leu Thr Val Phe Leu Asn Met Val Ala Glu Thr Met Pro

560

565

570

gct act tcc gat gcg gtg cca ttg tgg ata cgc atc gtg ttt ttg tgc 2138

Ala Thr Ser Asp Ala Val Pro Leu Trp Ile Arg Ile Val Phe Leu Cys

575

580

585

tgg ctg cca tgg ata ttg cga atg agt cgc cca gga cga ccg ctg atc 2186

Trp Leu Pro Trp Ile Leu Arg Met Ser Arg Pro Gly Arg Pro Leu Ile

590

595

600

605

cta gag ttc ccg acc acg ccc tgt tgc gac aca tcc tcc gag cgg aag 2234

Leu Glu Phe Pro Thr Thr Pro Cys Ser Asp Thr Ser Ser Glu Arg Lys

610

615

620

cac cag ata ctc tcc gac gtt gag ctg aaa gag cgc tgc tgc aaa tgc 2282

His Gln Ile Leu Ser Asp Val Glu Leu Lys Glu Arg Ser Ser Lys Ser

625

630

635

ctg ctg gcc aac gta cta gac atc gat gat gac ttc cgg cac aat tgt 2330

Leu Leu Ala Asn Val Leu Asp Ile Asp Asp Asp Phe Arg His Asn Cys

640

645

650

cgc ccc atg acg ccc ggc gga aca ctg cca cac aac ccg gct ttc tat 2378

Arg Pro Met Thr Pro Gly Gly Thr Leu Pro His Asn Pro Ala Phe Tyr

655

660

665

cgc acg gtt tat gga caa ggc gac gat ggc agc att ggg cca att ggc 2426

Arg Thr Val Tyr Gly Gln Gly Asp Asp Gly Ser Ile Gly Pro Ile Gly

670

675

680

685

agc acc cga atg ccg gat gcg gtc acc cat cat acg tgc atc aaa tca 2474

Ser Thr Arg Met Pro Asp Ala Val Thr His His Thr Cys Ile Lys Ser

690

695

700

tca act gaa tat gaa tta ggt tta atc tta aag gaa att cgc ttt ata 2522

Ser Thr Glu Tyr Glu Leu Gly Leu Ile Leu Lys Glu Ile Arg Phe Ile

705

710

715

act gat cag cta cgt aaa gat gac gag tgc aat gac att gcc aat gat 2570

Thr Asp Gln Leu Arg Lys Asp Asp Glu Cys Asn Asp Ile Ala Asn Asp

720

725

730

tgg aaa ttt gca gct atg gtc gtt gac aga ctg tgc ctt atc ata ttc 2618

Trp Lys Phe Ala Ala Met Val Val Asp Arg Leu Cys Leu Ile Ile Phe

735

740

745

aca atg ttc gca ata tta gcc aca ata gct gta cta cta tcg gca cca 2666

Thr Met Phe Ala Ile Leu Ala Thr Ile Ala Val Leu Leu Ser Ala Pro

750

755

760

765

cat att att gtc tcg tagccatattg ggcgaggtgg ttattgttat tggttttatt 2721

His Ile Ile Val Ser

770

ataaaatcaa ttgtttaatt attaaattaa taacgaaact ctttaagtaa attaaaacta 2781

aaaagacact aaaaaagcac aaaaaaatag gaaaatacat gataaaaccc atgaactaaa 2841

taatacatcc aagaaaaacc aaaacaaaaa aaaaaaaaaa aaaaa

2886

<210> 3

<211> 3701

<212> DNA

<213> *Heliothis virescens*

<220>

<221> CDS

<222> (335)..(1822)

<400> 3

ggcagagacc gctgccccac ggtcggcgc actccgctga acaacaatgc tcaaaaacac 60

gccgtgactc cacacacatc cctcggcgc agtaggcgat gtttgaggat cggacggcac 120

gcgtggccgt cggcgagcgg tcgtgaacaa gttgcataca tatgaaaacc gtaaaaagat 180

tgaattttaa gccgatcgtg ttcgatagat cctaatagag aagcgggagt gcggcgtttg 240

gtaggcgggg gtcgagtcgc gcggtcgggg gaaatggcgc ggcgcggggc ggcggcggcg 300

gcggcgcgcg gcgcggcggc gtcgcggcgc tgac atg ggc ggg cgg gcg cgc cgc 355

Met Gly Gly Arg Ala Arg Arg

1

5

tcg cac ttg gcg gcg ccc gcg ggc ctg ctg ctg ctg tgc ctg ctc 403

Ser His Leu Ala Ala Pro Ala Gly Leu Leu Leu Leu Cys Leu Leu

10

15

20

tgg ccg agg ggg gca cgc tgc ggg tac cac gag aag cgg cta ctg cac 451

Tyr Pro Arg Gly Ala Arg Cys Gly Tyr His Glu Lys Arg Leu Leu His

25

30

35

cac cta ttg gac cac tac aac gta ctg gag agg ccc gtc gtc aac gag 499

His Leu Leu Asp His Tyr Asn Val Leu Glu Arg Pro Val Val Asn Glu

40

45

50

55

agc gac ccg ctg cag ctc tcc ttc ggc ctc acg ctc atg cag atc atc 547

Ser Asp Pro Leu Gln Leu Ser Phe Gly Leu Thr Leu Met Gln Ile Ile

60

65

70

gac gtg gac gag aag aac cag ctt tta ata aca aac atc tgg cta aaa 595

Asp Val Asp Glu Lys Asn Gln Leu Leu Ile Thr Asn Ile Trp Leu Lys

75

80

85

cta gag tgg aat gat atg aac ttg agg tgg aac act tca gat ttc ggc 643

Leu Glu Trp Asn Asp Met Asn Leu Arg Trp Asn Thr Ser Asp Phe Gly

90

95

100

ggg gtc aaa gat tta aga gtg cca ccc cac aga cta tgg aaa cca gac 691

Gly Val Lys Asp Leu Arg Val Pro Pro His Arg Leu Trp Lys Pro Asp

105

110

115

gtc ctt atg tac aac agc gcg gac gaa ggg ttc gac agc acg tat cca 739

Val Leu Met Tyr Asn Ser Ala Asp Glu Gly Phe Asp Ser Thr Tyr Pro

120

125

130

135

acg aac gtg gtg gtg cgg aac aac ggc tcg tgt ctg tac gtg ccg ccc 787

Thr Asn Val Val Val Arg Asn Asn Gly Ser Cys Leu Tyr Val Pro Pro

140

145

150

ggc atc ttc aag agc acc tgc aag atc gac atc acc tgg ttc ccc ttc 835

Gly Ile Phe Lys Ser Thr Cys Lys Ile Asp Ile Thr Trp Phe Pro Phe

155 160 165
 gac gac caa cga tgc gag atg aag ttt ggc agc tgg act tat gat ggt 883
 Asp Asp Gln Arg Cys Glu Met Lys Phe Gly Ser Trp Thr Tyr Asp Gly
 170 175 180
 tat cag ttg gat cta caa cta cag gat gaa ggg ggc gga gat ata agc 931
 Tyr Gln Leu Asp Leu Gln Leu Gln Asp Glu Gly Gly Gly Asp Ile Ser
 185 190 195
 agt ttt gtc acg aat ggc gaa tgg gag tta ata gga gtc ccc ggc aag 979
 Ser Phe Val Thr Asn Gly Glu Trp Glu Leu Ile Gly Val Pro Gly Lys
 200 205 210 215
 cgc aac gag atc tac tac aac tgt tgt ccg gag cca tac atc gac atc 1027
 Arg Asn Glu Ile Tyr Tyr Asn Cys Cys Pro Glu Pro Tyr Ile Asp Ile
 220 225 230
 acg ttt gcg gtg gtg atc cgg agg aaa acg ctc tac tac ttc ttc aat 1075
 Thr Phe Ala Val Val Ile Arg Arg Lys Thr Leu Tyr Tyr Phe Phe Asn
 235 240 245
 ctg atc gtg ccc tgc gtg ctc atc gcc tcc atg gct cta ttg ggg ttc 1123
 Leu Ile Val Pro Cys Val Leu Ile Ala Ser Met Ala Leu Leu Gly Phe
 250 255 260
 acc ttg cct cca gac tcc gga gaa aag ttg tct tta ggt gtg acg ata 1171
 Thr Leu Pro Pro Asp Ser Gly Glu Lys Leu Ser Leu Gly Val Thr Ile
 265 270 275
 tta ctg tcg ttg acg gtg ttc ctc aac atg gtg gcg gag acg atg cca 1219

Leu Leu Ser Leu Thr Val Phe Leu Asn Met Val Ala Glu Thr Met Pro
 280 285 290 295

gcg acg tcg gac gcc gtg ccc ttg ctc ggc acc tac ttc aac tgc atc 1267
 Ala Thr Ser Asp Ala Val Pro Leu Leu Gly Thr Tyr Phe Asn Cys Ile
 300 305 310

atg ttc atg gtg gct tcc tcc gtc gtc tcc acc ata ctg atc ctc aac 1315
 Met Phe Met Val Ala Ser Ser Val Val Ser Thr Ile Leu Ile Leu Asn
 315 320 325

tac cac cac cgg cac gca gac act cac gaa atg agt gat tgg att cgt 1363
 Tyr His His Arg His Ala Asp Thr His Glu Met Ser Asp Trp Ile Arg
 330 335 340

tgc gtg ttc ctt tat tgg ctg ccg tgg gtg ctg cgc atg tca cgg ccc 1411
 Cys Val Phe Leu Tyr Trp Leu Pro Trp Val Leu Arg Met Ser Arg Pro
 345 350 355

ggc tcg gcg acg acg ccg ccg ccg gcg cgc gta cct ccg ccg ccg gac 1459
 Gly Ser Ala Thr Thr Pro Pro Pro Ala Arg Val Pro Pro Pro Pro Asp
 360 365 370 375

ctg gag ctg cgc gag cgc tcc tcc aag tcg ctc cta gcg aac gtg ctc 1507
 Leu Glu Leu Arg Glu Arg Ser Ser Lys Ser Leu Leu Ala Asn Val Leu
 380 385 390

gac atc gat gac gac ttc cgc cac ccg caa gcg cag cag ccg caa tgc 1555
 Asp Ile Asp Asp Asp Phe Arg His Pro Gln Ala Gln Gln Pro Gln Cys
 395 400 405

1507
 1459
 1363
 1315
 1267

tgc cga tac tac agg ggg ggt gag gag aat ggc gcg ggg ttg gcg gcg 1603
 Cys Arg Tyr Tyr Arg Gly Gly Glu Glu Asn Gly Ala Gly Leu Ala Ala

410

415

420

cac agt tgc ttc ggt gtc gac tac gag ctc tcc ctc att ctg aag gag 1651
 His Ser Cys Phe Gly Val Asp Tyr Glu Leu Ser Leu Ile Leu Lys Glu

425

430

435

att aga gtc atc aca gat cag atg cgc aag gac gac gaa gat gcg gac 1699
 Ile Arg Val Ile Thr Asp Gln Met Arg Lys Asp Asp Glu Asp Ala Asp

440

445

450

455

att tcg cgc gac tgg aag ttc gcc gcc atg gtc gtg gac aga ctg tgc 1747
 Ile Ser Arg Asp Trp Lys Phe Ala Ala Met Val Val Asp Arg Leu Cys

460

465

470

ctt att atc ttt acc ctg ttc aca atc atc gcc acg cta gcc gtg ctg 1795
 Leu Ile Ile Phe Thr Leu Phe Thr Ile Ile Ala Thr Leu Ala Val Leu

475

480

485

ctg tcc gcg cca cac atc atg gtg tcg tagcgacccg cccgcttgcg 1842
 Leu Ser Ala Pro His Ile Met Val Ser

490

495

gatacgcatg cgaaaagttc tgtgataccg cgaatatttg ttaagttgtg atgagcgaag 1902

tggcgcgac ggtgacgccg cggcgtcgga gttgccgccg cctgcctcgc cgccccgcgc 1962

cccctgtaga cataagttac cgctgactgc caaccctgta cgttcaacaa ataactgccc 2022

atccgactaa cgtcttttat ccccttgaaa aattcagcga ttgtgtaccc ctttcttcca 2082

cattttataa gatatacaac attttataag tacaagaagt tgtaacatga accgggttttt 2982

cgttacatag aggggtataac acaaaggtgc ctacatattg acagatgcga agcacgatca 3042

gttgataagc acaggtacac tatatcctga catccgacag tcctgccgct cgtctgccac 3102

actcggaaac attcgacagt tcagtttact gtcgcgcat catcgattgt taagtttggt 3162

gttctaactc atcgattca tttcattcaa aaacattgta aacctctcaa ggggaaaacg 3222

tgttgtaaac agtgagagtg cgcgggtaca accgacacgc gaatgtaccc tcgcaaggct 3282

cctgtaatgt tttcctcttc cgaggtgttg ctgagagtaa tcttagacgg tccgatggaa 3342

gttgccgacc ggatatgatt acaagtcaat gtttttaagt catccgttta tttattgtta 3402

tatcttctta ccattcgcta gaggttggtg gacgaccgg acggtggggc ccgcaaccgc 3462

cacacgcggg gttccatctt tgtattagat ggaagttgtg cggcatctct ccgtcggcaa 3522

tgggacaacc cgttggtccc aacatttggt caattgtag ggttaactct gaattgcact 3582

ttgtttatta aatataaacg aatgaaacaa aaaaaaaaaa aaaaaactcg agagtacttc 3642

tagagcggcc gcgggcccac cgattttcca cccgggtggg gtaccargta agtgtaccc 3701

CGTACATAG AGGGGTATAAC ACAAGGTGC CTACATATTG ACAGATGC GAAGCACGATCA 3042
GTTGATAAGC ACAGGTACAC TATATCCTGA CATCCGACAG TCCTGCCGCT CGTCTGCCAC 3102
ACTCGGAAAC ATTCGACAGT TCAGTTTACT GTCGCGCAT CATCGATTGT TAAGTTTGTT 3162
GTTCTAACTC ATCGATTCA TTTCAATCAA AAACATTGTA AACCTCTCAA GGGGAAAACG 3222
TGTTGTTAAAC AGTGAGAGTG CGCGGGTACA ACCGACACGC GAATGTACCC TCGCAAGGCT 3282
CCTGTAATGT TTTCTCTTC CGAGGTGTTG CTGAGAGTAA TCTTAGACGG TCCGATGGAA 3342
GTTGCCGACC GGATATGATT ACAAGTCAAT GTTTTAAAGT CATCCGTTTA TTTATTGTTA 3402
TATCTTCTTA CCATTGCTA GAGGTGTGTG GACGACCGG ACGGTGGGGC CCGCAACCCG 3462
CACACGCGGG GTTCCATCTT TGTATTAGAT GGAAGTTGTG CGGCATCTCT CCGTCGGCAA 3522
TGGGACAACC CGTTGTCCC AACATTGTT CAATTGTAG GGTAACTCT GAATTGCCT 3582
TTGTTTATTA AATATAAACG AATGAAACAA AAAAAAAAAA AAAAACTCG AGAGTACTTC 3642
TAGAGCGGCC GCGGGCCAC CGATTTTCCA CCCGGGTGGG GTACCARGTA AGTGTACCC 3701

<210> 5

<211> 3109

<212> DNA

<213> *Heliothis virescens*

<220>

<221> CDS

<222> (95) .. (1597)

<400> 5

ggcacgagcc ggccgcacgt tgtcccaggc cgcatagacg cgccggcggtg ctagcgcagc 60

gtgcgcgggt gtggtatgcc cgcgcgtcgc cgt atg gcc cct atg ttg gcg gcc 115

Met Ala Pro Met Leu Ala Ala

1

5

ttg gcg ctg ctg gct ttg ctg ccc gta tcg gag caa ggt cct cac gag 163

Leu Ala Leu Leu Ala Leu Leu Pro Val Ser Glu Gln Gly Pro His Glu

10

15

20

aag aga ctc ctg aac gcg ttg ctg gcg aac tac aac acc ctg gag cga 211

Lys Arg Leu Leu Asn Ala Leu Leu Ala Asn Tyr Asn Thr Leu Glu Arg

25

30

35

ccg gtg gcc aac gag agc gaa ccg cta gag gtc agg ttc ggc ttg acc 259

Pro Val Ala Asn Glu Ser Glu Pro Leu Glu Val Arg Phe Gly Leu Thr

40

45

50

55

ttg cag caa atc att gac gtg gac gag aag aat caa cta ctt ata acc 307

Leu Gln Gln Ile Ile Asp Val Asp Glu Lys Asn Gln Leu Leu Ile Thr

60

65

70

aat ata tgg ctg tct ttg gag tgg aat gac tac aac ctg agg tgg aac 355

Asn Ile Trp Leu Ser Leu Glu Trp Asn Asp Tyr Asn Leu Arg Trp Asn

75

80

85

gac agc gag tat ggc ggg gtc aag gac ctc agg atc acg ccc aac aag 403

Asp Ser Glu Tyr Gly Gly Val Lys Asp Leu Arg Ile Thr Pro Asn Lys

90

95

100

ttg tgg aag ccg gac gtc ctt atg tat aat agt gct gac gag ggt ttt 451

Leu Trp Lys Pro Asp Val Leu Met Tyr Asn Ser Ala Asp Glu Gly Phe

105

110

115

gac ggg acc tac cag acc aac gtg gtg gtc aga agc ggc ggc agt tgc 499

Asp Gly Thr Tyr Gln Thr Asn Val Val Val Arg Ser Gly Gly Ser Cys

120

125

130

135

ctg tac gtg cca cct ggc ata ttc aag agc aca tgc aag atg gac atc 547

Leu Tyr Val Pro Pro Gly Ile Phe Lys Ser Thr Cys Lys Met Asp Ile

140

145

150

gcg tgg ttc ccc ttc gac gac caa cac tgt gat atg aag ttc ggt agc 595

Ala Trp Phe Pro Phe Asp Asp Gln His Cys Asp Met Lys Phe Gly Ser

155

160

165

tgg aca tat gac ggc aat cag ttg gat ctg gtg cta aaa gat gag gca 643

Trp Thr Tyr Asp Gly Asn Gln Leu Asp Leu Val Leu Lys Asp Glu Ala

170

175

180

ggc ggc gat cta tcg gac ttc ata aca aat ggg gag tgg tat cta ata 691
 Gly Gly Asp Leu Ser Asp Phe Ile Thr Asn Gly Glu Trp Tyr Leu Ile

185

190

195

gga atg cca ggc aaa aag aac aca ata aca tac gcg tgc tgc ccc gag 739
 Gly Met Pro Gly Lys Lys Asn Thr Ile Thr Tyr Ala Cys Cys Pro Glu

200

205

210

215

ccc tac gtg gac gtc acc ttc acc atc atg ata aga aga cga acc ttg 787
 Pro Tyr Val Asp Val Thr Phe Thr Ile Met Ile Arg Arg Arg Thr Leu

220

225

230

tac tac ttc ttc aac ctg atc gtc ccg tgc gtg ctg atc tca tcg atg 835
 Tyr Tyr Phe Phe Asn Leu Ile Val Pro Cys Val Leu Ile Ser Ser Met

235

240

245

gca ctc ctc ggc ttc aca ctg cca cca gac tcc gga gag aaa ctc aca 883
 Ala Leu Leu Gly Phe Thr Leu Pro Pro Asp Ser Gly Glu Lys Leu Thr

250

255

260

ctt gga gtc act att ctt cta tcg ctg acg gtg ttc ctc aac ctg gta 931
 Leu Gly Val Thr Ile Leu Leu Ser Leu Thr Val Phe Leu Asn Leu Val

265

270

275

gcc gag acc ctg cca cag gtc tcc gac gct atc ccc ctg tta ggg acg 979
 Ala Glu Thr Leu Pro Gln Val Ser Asp Ala Ile Pro Leu Leu Gly Thr

280

285

290

295

tac ttc aat tgc atc atg ttc atg gta gcg tcg tct gtg gta ctg act 1027
 Tyr Phe Asn Cys Ile Met Phe Met Val Ala Ser Ser Val Val Leu Thr

300

305

310

gtg gtg gta ctc aat tac cac cat cga aca gct gat ata cat gaa atg 1075

Val Val Val Leu Asn Tyr His His Arg Thr Ala Asp Ile His Glu Met

315

320

325

cca cag tgg ata aaa tca gta ttc cta caa tgg ttg cca tgg ata ctg 1123

Pro Gln Trp Ile Lys Ser Val Phe Leu Gln Trp Leu Pro Trp Ile Leu

330

335

340

cga atg tcg agg cca ggg aag aag atc acc agg aag act ata atg atg 1171

Arg Met Ser Arg Pro Gly Lys Lys Ile Thr Arg Lys Thr Ile Met Met

345

350

355

aac acg agg atg agg gag ctg gaa ctg aag gag agg tcg tcg aag tcc 1219

Asn Thr Arg Met Arg Glu Leu Glu Leu Lys Glu Arg Ser Ser Lys Ser

360

365

370

375

ttg ctg gcg aat gtt cta gat att gat gat gac ttc aga cac ggc cct 1267

Leu Leu Ala Asn Val Leu Asp Ile Asp Asp Asp Phe Arg His Gly Pro

380

385

390

ccg cct cct aac agt act gcc tcg acc ggg aat ttg gga cct ggg tgc 1315

Pro Pro Pro Asn Ser Thr Ala Ser Thr Gly Asn Leu Gly Pro Gly Cys

395

400

405

tca ata ttc cgc acg gat ttc cgt cgg tcg ttc gtc cgt ccg tcc acg 1363

Ser Ile Phe Arg Thr Asp Phe Arg Arg Ser Phe Val Arg Pro Ser Thr

410

415

420

atg gaa gac gtg ggc ggc ggg ctg ggt agc cac cat cgc gag ctg cac 1411

Met Glu Asp Val Gly Gly Gly Leu Gly Ser His His Arg Glu Leu His

435

Leu Ile Leu Arg Glu Leu Gln Phe Ile Thr Ala Arg Met Lys Lys Ala

455

Asp Glu Glu Ala Glu Leu Ile Ser Asp Trp Lys Phe Ala Ala Met Val

470

Val Asp Arg Phe Cys Leu Phe Val Phe Thr Leu Phe Thr Ile Ile Ala

485

Thr Val Ala Val Leu Leu Ser Ala Pro His Ile Ile Val Gln

500

ccgatttgta attataattg/ataatgtaat taaattaaat acgtggttga aacgcacacg 1717

gagtgttgat g~~g~~aaatttta gccggcgcaa ggagtttcgt gaaggctctgt atatattttt 1837

tcttattggt gtatattgta tcgttgttca tgttttcttt caggaagtga gctttgtact 1897

gtttg/ttct tcgatggcag gtgcacttca gttcaggctg aaatttccat taacatttat 1957

ttaaacaat gtgatgttga ctaggatgtt atacagataa atgttgacgt gtataatttg 2017

ttaaaataaa caatattaat tactattact aaacgatatt ataaacgaag tactaacgag 2077
ggttacttta atgggaagaa cgctaagctg gcacagagtt gcattaattt gaaaaaagaa 2137
attacggaaa aaagtttatt gaaaattgaa ctttttgga ggaaagtaac gtttgatcaa 2197
aaaagtttgt aaaacgaaag ttcggttctg cgccaatact ggaattaaaa ttctcgtaaa 2257
tattagggaa aagaaggtcc tttaaaacaa aagatttgaa ccggcatcct tttacaagt 2317
aatgagggat cacagatgat gacaaaaaac cttagggat ataagtaatg tacataatgg 2377
atcaaatac ggtagagtca agaatagtta acgatttaag attattccat tcgatattaa 2437
aattcgatta gcgattgtcg ctgcgtctac ttgatacat atcgatttga atcgatattg 2497
tataaattta gatagatcgg acattagtaa tgagtatgga cgttttaatt tttaaaaaag 2557
aatgtactac gaagattaaa tccaggaatt gttaaacagt tatggaattg ataagaaatc 2617
aacaattaat acggaccaa aggtagacta ggtgtagcat caggagattg aattaaaaca 2677
taaattagga ccgacttaa tggaacttgc gagtgtattg ataacttttt aatttaaaaa 2737
ctcattgtcg attaaatgga gaataacttt tgatctctcg tatcgataaa tgctcactta 2797
actatcgata gcgtaatat ataactgtta gtatatcgat atgggagtaa gtcactagca 2857
tcagaaatag tcattaatta ggaatcggtt tgtgttaatg ttatgcttag cgaaaatatt 2917

acaatgctgt tgatatcact aaccatcacg taaccatatt gataaaatgt aaatacagaa 2977

tattgcggtg tgtatttgta tataaatTTT agaaaaaaaa aaaaaaaaaa aactcgagag 3037

tacttctaga gcggccgcgg gcccatcgat tttccaccg ggtgggggtac caggtaagtg 3097

taccaattc gc

3109